



Centre Interuniversitaire sur le Risque,
les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **12-05**

On the Value of Improved Informativeness

Pierre Chaigneau

Février/February 2012

Chaigneau: Department of Finance, HEC Montréal, 3000 chemin de la côte Sainte-Catherine, Montréal, QC, Canada
H3T 2A7. Tel.: +1 514 340-5661; and CIRPÉE
pierre.chaigneau@hec.ca

Abstract:

One of the main predictions of principal-agent theory, the “informativeness principle”, is often violated in practice. We propose an explanation that emphasizes the role played by the change in the form of the optimal contract that follows an improvement in informativeness. We show that the overall gains from a less noisy performance measure emanate from two sources: the direct effect of a change in the volatility of the performance measure, and the effect of the induced change on the form of the optimal compensation contract. We emphasize that the direct effect can either largely underestimate or largely overestimate the overall gains from improved informativeness, and we show that these gains can even be nil in some instances.

Keywords: Executive compensation, pay-for-luck, principal-agent model, relative performance evaluation, stock-options

JEL Classification: D86, J33

“One should try to use the most informative or precise measurement of the agent’s economic activity, or what is called in statistics a summary or sufficient statistic. This result underlies much of the thinking about managerial compensation, the quest for good metrics to reward employees or division managers.” Jean Tirole, *The Theory of Corporate Finance*

As emphasized in this introductory quote, the design of performance measures should in theory rely on the so-called “informativeness principle”, which is arguably one of most important results of principal-agent theory. It is supported by results in Shavell (1979), Holmstrom (1979), Gjesdal (1982), Grossman and Hart (1983), Kim (1995), and is reviewed in the textbooks of Salanié (1997), Laffont and Martimort (2002), and Bolton and Dewatripont (2005).¹ However, numerous violations of this principle have been identified in the empirical literature. They include the absence of relative performance evaluation (Aggarwal and Samwick (1999), Murphy (1999)), pay-for-luck (Bertrand and Mullainathan (2001)), i.e., a positive sensitivity of pay to exogenous and contractible shocks to performance, and grants of stock-options to non-CEOs (Core and Guay (2001), Oyer and Schaefer (2005)). This paper proposes a novel, albeit partial, explanation based on constraints on contracting.

Many papers refer to the “informativeness principle” to argue that any contract which does not filter out exogenous and contractible shocks to performance is suboptimal, in a standard model of moral hazard. However, there is little research on the value of improved informativeness. This is an important question: if it is sufficiently low in certain contexts, then the gains may not outweigh the direct and possibly indirect costs of altering the compensation system. In any case, informativeness could then, in those instances at least, be treated as a side-issue.

A recent paper by Dittmann, Maug and Spalt (2011) assesses the value of increased infor-

¹Shavell (1979) shows that additional information on the agent’s effort has positive value. Gjesdal (1982) and Grossman and Hart (1983) show that if the information structure A is sufficient for the information structure B in the sense of Blackwell, then A is associated with a lower agency cost than B . Holmstrom (1979) shows that any signal which is informative about the agent’s action will be included in the compensation contract. Lastly, Kim (1995) shows that the information structure A is more efficient than B if the cumulative distribution function of the likelihood ratio under A is a mean-preserving spread of the one under B .

mativeness in a principal-agent model of CEO pay. Their calibration approach enables them to assign dollar values to changes in informativeness for a sample of U.S. CEOs. A limitation of their approach is that they take the form of the compensation contract as given.² In this paper, we emphasize that it is important to assess not only the direct effect of increased informativeness, holding all else constant, but also its indirect effect on the form of the optimal contract.

To incorporate this indirect effect, we derive endogenously the form of the optimal contract in a stylized principal-agent model of effort choice. We use a setting similar to Innes (1990), which is now standard in many fields, including corporate finance (e.g. Tirole (2006)). In particular, the principal and the agent are risk neutral, the agent is protected by limited liability, and the sensitivity of pay to performance is bounded from above.³ As in Innes, we first show that call-options-based compensation is optimal, for any level of informativeness of the performance measure.⁴

We then assess the gains achievable by altering the informativeness of the performance measure. In practice, this can be achieved by filtering out exogenous shocks to performance, for example. Once the change in the shape of the compensation contract (which we refer to as the “indirect effect”) is accounted for, we show both theoretically and numerically that these gains are not only minimal in some instances, but also much smaller than with a partial equilibrium approach which omits this indirect effect. This is the case for low-powered

²Their approach consists in calibrating a CRRA-lognormal model to a sample of CEOs. Since this model does not generate observed compensation contracts (Dittmann and Maug (2007)), they cannot endogeneize the contract. Instead, following a change in the informativeness or “indexing” of stock-options, they postulate that the fixed salary and the number of restricted shares adjusts to satisfy the participation and the incentive constraints.

³The model of Innes is also presented in the textbooks of Bolton and Dewatripont (2005, section 4.6.2) and Tirole (2006, section 3.6). Even though the informativeness principle was originally derived in a model with a risk averse agent, it extends to settings with a risk neutral agent protected by limited liability. This point is made clear for example in Tirole (2006), in a setting without any constraint on contracting except for limited liability – see section 3.2.6., page 123 of the first edition, especially footnote 23.

⁴By contrast, a standard CRRA-lognormal principal-agent model cannot explain the use of stock-options (Hall and Murphy (2002), Dittmann and Maug (2007)). This being said, Hemmer, Kim, and Verrecchia (2000) can generate convex compensation contracts in a HARA-gamma framework, while Dewatripont, Legros, and Matthews (2003) show that stock-options (or equivalently a levered firm with equity held by the agent and debt held by outside investors) can be optimal in a model with a risk averse agent and renegotiation after effort is observed.

incentives when the performance measure is normally distributed, for example, in which case the overall gains are even nil in the limit, as incentives are extremely low-powered. Thus, to the extent that improving the informativeness of the performance measure is (even slightly) costly, there will be instances when it is suboptimal to do so. By showing that pay-for-luck (Bertrand and Mullainathan (2001)) and stock-options can be optimal even when the cost of improving the informativeness of the performance measure is very low, the model can contribute to explain both the compensation base and the main form of compensation of most executives.⁵

This result is explained by the presence of constraints on contracting. Because both the limited liability constraint and the upper bound constraint on the pay-performance sensitivity are binding in equilibrium, and because a call option contract is optimal for any given variance of the performance measure (as in Innes (1990)), the exercise price of the compensation contract adjusts for incentive compatibility. In particular, a change in the variance of the performance measure induces a change in the exercise price. We show that, for sufficiently high-powered incentives and a normally distributed performance measure, the exercise price increases following a reduction in the variance of the performance measure, which further reduces the agency rent. In this case, improving the informativeness of the performance measure is beneficial because of the well-known direct effect on the agency rent *and* because of this indirect effect on the exercise price. However, for sufficiently low-powered incentives, the exercise price diminishes following a reduction in the variance of the performance measure, which contributes to increase the expected payout from the contract and therefore the agency rent. In this case, the adverse change in the exercise price mitigates the gains from improved informativeness.

Other explanations for pay-for-luck have been proposed in the literature. Oyer (2004) shows that pay-for-luck may be optimal if the value of employees' outside options vary with

⁵Murphy (1999) shows that stock-options are the major component of CEO pay. Frydman and Jenter (2010) corroborate this finding for the 2000-2005 period. In addition, many firms grant stock-options not only to their CEOs, but also to most of their employees (Oyer and Schaefer (2005)).

economic conditions and if re-contracting is costly. Garvey and Milbourn (2006), Danthine and Donaldson (2008), Axelson and Baliga (2009), and Gopalan, Milbourn and Song (2010) also propose explanations for pay-for-luck (this literature is reviewed in Edmans and Gabaix (2009)). It may also be preferable to base compensation on measures of output rather than measures of input when the agent has some private information on the production technology (Raith (2008)).

Section 1 presents the model. Section 2 derives the form of the optimal contract as a function of the performance measure, for a given variance of the performance measure. Given this form of the optimal contract, section 3 derives the gains from a reduction in the variance of the performance measure. Section 4 concludes. All proofs are in the Appendix.

1 The model

We consider a standard stylized principal-agent model of effort choice in the spirit of Innes (1990), in which a firm is owned by a principal and managed by an agent. The timeline is the following. At time -1 , the principal offers a compensation contract W to the agent. At time 0, the agent chooses her level of effort e . Effort affects the probability distribution of contractible profits \tilde{x} , which are realized at time 1. The problem of the principal is to design the compensation contract to maximize the expected value of profits \tilde{x} net of the expected cost of compensation.⁶ The discount factor is zero.

The agent is risk neutral. She maximizes her expected payment net of effort cost. Her reservation wage is normalized at zero. At time 0, she exerts either low effort $e = 0$ at no cost, or high effort $e = \bar{e}$ at cost C (in monetary terms). The choice of e is unobservable.

⁶Thus, we adopt the approach which is most standard in the literature, but in doing so we differ from Innes (1990). Mathematically, in this paper the objective function of the principal is maximized subject to the incentive constraint and a minimum value for the objective function of the agent, whereas in the Innes paper the objective function of the agent (the entrepreneur) is maximized subject to the same incentive constraint and a minimum value for the objective function of the principal (the investor). As explained in the footnote 2 of the Innes paper, these two optimization problems yield the same optimal contracts. Indeed, the set of all feasible efficient contracts is obtained by adjusting the minimum value for the objective function of the agent in this paper, or by adjusting the minimum value for the objective function of the principal in Innes' paper.

We assume that profits may be written as $\tilde{x} = e + \tilde{v} + \tilde{\epsilon}$. The random variable \tilde{v} is unobservable, and normally distributed, with mean zero and variance $\underline{\sigma}^2$. The random variable $\tilde{\epsilon}$ is normally distributed, with mean zero and variance $\bar{\sigma}^2 - \underline{\sigma}^2$. The two random variables \tilde{v} and $\tilde{\epsilon}$ are independent. The normal distribution facilitates tractability and suits our purposes: we need a probability distribution which satisfies the monotone likelihood ratio property, and for which mean-preserving spreads can be generated by altering one parameter (the logistic distribution shares these properties, but it is not as tractable for our purposes). For these reasons and others, the normal distribution is often used in the related literature. We discuss the robustness of our results to this normality assumption, and its limitations, in section 3.2. Moreover, since we focus on informativeness improvements at the margin, we assume that $\bar{\sigma} - \underline{\sigma}$ is arbitrarily small. Notice that expected profits are equal to the agent's effort – this is without loss of generality: expected profits could be equal to the agent's effort plus a constant.

Crucially, the agent's compensation is not contingent on \tilde{x} per se, but on a measure of performance $\tilde{\pi}$. This measure of performance may or may not include $\tilde{\epsilon}$: we assume that a technology allows the firm to filter out the shock $\tilde{\epsilon}$ if it pays the cost κ at time -1 .⁷ More precisely, we let the dummy variable δ be equal to zero if the firm filters the shock $\tilde{\epsilon}$ out of the measure of performance $\tilde{\pi}$, and equal to one otherwise, so that the measure of performance writes as:

$$\tilde{\pi} = e + \tilde{v} + \delta \tilde{\epsilon} \tag{1}$$

Thus, for $\delta = 1$, the shock $\tilde{\epsilon}$ is not filtered out of the measure of performance, and the agent is “paid-for-luck”. On the contrary, for $\delta = 0$, the shock $\tilde{\epsilon}$ is filtered out of the measure of performance, whose informativeness is therefore maximized.

Roughly speaking, we study the value of informativeness in terms of a mean-preserving reduction in the dispersion of probability mass. This is apparent in the decomposition of \tilde{x}

⁷Several interpretations are possible. The firm may make the agent's compensation contingent upon an index calculated by a third party. Alternatively, the firm may use risk management techniques to reduce the volatility of profits, in which case $\tilde{\pi}$ with $\delta = 1$ represents the profits without risk management, and $\tilde{\pi}$ with $\delta = 0$ represents the profits with risk management.

into the agent's effort, and two independent white noises. Since \tilde{x} is additive in these three terms, filtering out the white noise $\tilde{\epsilon}$ leads to a mean-preserving change in the distribution of the performance measure $\tilde{\pi}$ in the sense of second-order stochastic dominance. We refer to these changes as improvements in informativeness.

Given $\delta \in \{0, 1\}$, let $\sigma^2(\delta)$ denote the variance of $\tilde{\pi}$, so that $\sigma^2(\delta) = \underline{\sigma}^2 + \delta^2(\bar{\sigma}^2 - \underline{\sigma}^2)$, and $\sigma(\delta) \in \{\underline{\sigma}, \bar{\sigma}\}$. Notice that choosing $\delta \in \{0, 1\}$ is equivalent to choosing $\sigma \in \{\underline{\sigma}, \bar{\sigma}\}$. To alleviate notations, we will write σ instead $\sigma(\delta)$ in what follows. We let φ and Φ denote the p.d.f. and the c.d.f. of the standard normal variable, respectively, and ψ denote the p.d.f. of $\tilde{\pi}$, for a given σ . The likelihood ratio is

$$LR(i) \equiv \frac{\psi(\pi = i|e = \bar{e}) - \psi(\pi = i|e = 0)}{\psi(\pi = i|e = 0)} = \frac{\varphi\left(\frac{i - \bar{e}}{\sigma}\right) - \varphi\left(\frac{i}{\sigma}\right)}{\varphi\left(\frac{i}{\sigma}\right)} \quad (2)$$

It is strictly increasing in i , so that the monotone likelihood ratio property is satisfied. For any given $\sigma \in [\underline{\sigma}, \bar{\sigma}]$, let \hat{Z} be implicitly defined by

$$\varphi\left(\frac{\hat{Z} - \bar{e}}{\sigma}\right) \equiv \varphi\left(\frac{\hat{Z}}{\sigma}\right) \quad (3)$$

which implies that $\hat{Z} = \frac{\bar{e}}{2}$. Note that \hat{Z} is the exercise price such that the likelihood ratio in (2) is zero when evaluated at $i = \hat{Z}$: \hat{Z} is the least informative performance. Since the likelihood ratio is strictly increasing, $LR(i) > 0$ if and only if $i > \hat{Z}$.

The contract offered by the principal to the agent specifies the payment to the agent $W(\pi)$ as a function of the realization of the measure of performance $\tilde{\pi}$. We follow Innes (1990) in making two important assumptions which restrict the set of feasible contracts. First, we assume that the agent is protected by limited liability and cannot earn more than π , so that $0 \leq W(\pi) \leq \max\{\pi, 0\}$, for any value of π . Second, we assume that the pay-performance-sensitivity is bounded from above: $W'(\pi) \leq b$ for all π , where $b \in (0, 1]$. Innes (1990) sets $b = 1$, and justifies this constraint on two grounds. First, if it did not hold on some interval,

then the agent could borrow on her own account to artificially increase the value of π on this interval, thus undoing the contract on this interval. Second, the principal would have incentives to exercise his control rights to “burn profits” or sabotage the firm on this interval, since for any given increase of the performance measure on this interval, payments to the agent increase even more. This would also undo the contract on the interval in question.

Given a contract $W(\pi)$ and a level of effort e , the expected payment to the agent is

$$E[W(\tilde{\pi})|e] = \int_{-\infty}^{\infty} W(\pi)\psi(\pi|e)d\pi$$

For a contract which is such that the agent optimally chooses $e = \bar{e}$, we define the agent’s agency rent as her expected payment in equilibrium net of effort cost:

$$AR \equiv E[W(\tilde{\pi})|e = \bar{e}] - C \quad (4)$$

If $E[\tilde{\pi}|e = \bar{e}] - E[\tilde{\pi}|e = 0] > C$, then the first-best is defined by $e = \bar{e}$ and $E[W(\tilde{\pi})|e = \bar{e}] = C$: the agent is only compensated for her cost of effort to guarantee her participation, and she does not derive any agency rent.

We assume that C is sufficiently low for the implementation of $e = \bar{e}$ rather than $e = 0$ to be second-best optimal (the problem would be trivial otherwise, with an optimal contract such that $W(\pi) = 0$ for all π). At the second-best, the problem of the principal is therefore to elicit $e = \bar{e}$ while minimizing the expected cost of compensation in equilibrium, which is equivalent to minimizing the agency rent in (4) net of the cost κ of setting $\delta = 0$ if incurred. The incentive constraint, which guarantees that the agent chooses $e = \bar{e}$ rather than $e = 0$, is

$$E[W(\tilde{\pi})|e = \bar{e}] \geq E[W(\tilde{\pi})|e = 0] + C \quad (5)$$

A contract which satisfies this constraint is said to be incentive-compatible. The incentive constraint will be binding at the optimum (suppose that it is not: then it is possible to

reduce $E[W(\tilde{\pi})|e = \bar{e}]$ without violating any constraint, which diminishes the agency rent in (4)), which implies that a constant wedge of C must be maintained between the expected payment conditional on high effort and the expected payment conditional on low effort. At the optimum, the agency rent defined in (4) is therefore:

$$AR = E[W(\tilde{\pi})|e = 0] \quad (6)$$

If $E[W(\tilde{\pi})|e = 0]$ is strictly positive, then the expected payment to the agent conditional on high effort will need to exceed the level required to compensate her for the cost of effort C , in order to provide adequate incentives for high rather than low effort. This in turn generates a positive agency rent. Finally, to guarantee that an incentive-compatible optimal contract exists, we assume for reasons which will become clear later that, for any $\sigma \in \{\underline{\sigma}, \bar{\sigma}\}$,

$$\int_0^\infty \pi(\psi(\pi|e = \bar{e}) - \psi(\pi|e = 0))d\pi > \frac{C}{b} \quad (7)$$

In summary, the principal chooses the volatility of the performance measure, and the functional form relating the agent's pay to her performance measure. First, the principal chooses $\delta \in \{0, 1\}$, which is equivalent to choosing the standard deviation σ of the performance measure $\tilde{\pi}$ in $\{\underline{\sigma}, \bar{\sigma}\}$, as already noted. Second, the principal chooses the form of the compensation contract $W(\pi)$ subject to the aforementioned constraints. His problem may thus be written as

$$\min_{W(\pi), \delta} E[W(\tilde{\pi})|e = 0] - (1 - \delta)\kappa \quad \text{s.t.} \quad E[W(\tilde{\pi})|e = \bar{e}] = E[W(\tilde{\pi})|e = 0] + C$$

$$\text{and} \quad 0 \leq W(\pi) \leq \max\{\pi, 0\} \text{ for all } \pi \quad \text{and} \quad W'(\pi) \leq b \text{ for all } \pi \quad \text{and} \quad \delta \in \{0, 1\}$$

2 The form of the optimal contract

The model described in the preceding section is very similar to the one in Innes (1990). Differences include a reversal of the roles of principal and agent (as discussed in footnote 5), a choice between two levels of effort only to simplify the analysis, and the assumption that profits are normally distributed (Innes only requires the probability distribution to satisfy the monotone likelihood ratio property).⁸ As expected, the result on the form of the optimal contract that we derive in this section is roughly the same as in Innes (1990).

In this section, we take the variance of the performance measure σ^2 as given. The optimal contract therefore solves the following problem:

$$\min_{W(\pi)} E[W(\tilde{\pi})|e = 0] \quad (8)$$

subject to

$$E[W(\tilde{\pi})|e = \bar{e}] - C = E[W(\tilde{\pi})|e = 0] \quad (9)$$

$$W(\pi) \geq 0 \quad \text{for all } \pi \quad (10)$$

$$W(\pi) \leq \max\{\pi, 0\} \quad \text{for all } \pi \quad (11)$$

$$W'(\pi) \leq b \quad \text{for all } \pi \quad (12)$$

If the optimal contract exists, then it takes the following form:

Proposition 1: The optimal contract is characterized by $W(\pi) = 0$ for all $\pi < Z$, and $W'(\pi) = b$ for all $\pi > Z$, for a given positive Z .

⁸One consequence of these differences is that deviations from the first-best take different forms. In Innes (1990), the degree of inefficiency is measured by the difference between the first-best level of effort and the second-best level of effort to which the agent can commit. Hence the importance of allowing for a continuum of effort levels in the Innes paper. In this paper, the objective is instead to minimize the agency rent received by the agent.

The two forces that drive this result are the constraints on contracting and the increasing likelihood ratio of the normal distribution. The intuition is as in Innes (1990). The absolute value of the likelihood ratio is highest at the tails of the distribution of $\tilde{\pi}$, which means that performances are most informative about the level of effort chosen by the agent at the tails of the distribution. In addition, the left tail cannot be used for incentive purposes because of the limited liability constraint, so that incentives are concentrated on the right tail. With an upper-bound on the sensitivity of pay to performance, the optimal contract involves call options on $\tilde{\pi}$ with a slope that binds this upper-bound constraint. This maximizes the likelihood that positive payments are received by an agent who exerted high effort rather than an agent who exerted low effort, which minimizes the agency rent.

In addition, we know the following about the exercise price Z of the call options.

Lemma 1: There is one and only one positive Z which satisfies the incentive constraint in (5) as an equality.

We denote the value of Z which satisfies the incentive constraint as an equality for a given value of σ by $Z(\sigma)$. As established in section 1, the exercise price $Z(\sigma)$ of the optimal contract is set to bind the incentive constraint, so that it solves the following equation:

$$\int_{Z(\sigma)}^{\infty} b(\pi - Z(\sigma))(\psi(\pi|e = \bar{e}) - \psi(\pi|e = 0))d\pi = C \quad (13)$$

Corollary 1: $Z(\sigma)$ is strictly decreasing in C .

The higher the cost of effort C , the higher-powered incentives must be, and the lower $Z(\sigma)$ is. This result is important, because the value of $Z(\sigma)$ will play a crucial role in the next section.

To summarize, the optimal contract is

$$W(\pi) = \max\{0, b(\pi - Z(\sigma))\} \quad (14)$$

This formula is the payoff of a quantity b of call options on π with exercise price $Z(\sigma)$. It is noteworthy that this result holds for any variance σ^2 of the probability distribution of $\tilde{\pi}$.

3 The value of increased informativeness

3.1 Theoretical results

We now determine the gains from increased informativeness. More precisely, we relate the agency rent to the variance σ^2 of the performance measure $\tilde{\pi}$.

Since the contract in (14) is optimal and satisfies the constraints (9), (10), (11), and (12) for any $\sigma \in \{\underline{\sigma}, \bar{\sigma}\}$, the problem of the principal is to choose σ in order to minimize the agency rent net of the cost κ of setting $\delta = 0$ with a compensation contract as in (14), where $Z(\sigma)$ solves (13). This optimization problem may be written as

$$\min_{\delta} \int_{Z(\sigma)}^{\infty} b(\pi - Z(\sigma)) \psi(\pi | e = 0) d\pi - (1 - \delta)\kappa \quad \text{s.t. (13) and } \delta \in \{0, 1\} \quad (15)$$

We first establish three preliminary results.

Lemma 2: $\frac{dZ(\sigma)}{d\sigma} > 0$ if and only if $Z > \hat{Z}$.

Lemma 3: For a given $\sigma \in \{\underline{\sigma}, \bar{\sigma}\}$, $\frac{dAR(\sigma)}{d\sigma} > 0$ if and only if

$$\frac{\varphi\left(\frac{Z(\sigma) - \bar{e}}{\sigma}\right) - \varphi\left(\frac{Z(\sigma)}{\sigma}\right)}{\varphi\left(\frac{Z(\sigma)}{\sigma}\right)} < \frac{\Phi\left(\frac{Z(\sigma)}{\sigma}\right) - \Phi\left(\frac{Z(\sigma) - \bar{e}}{\sigma}\right)}{1 - \Phi\left(\frac{Z(\sigma)}{\sigma}\right)} \quad (16)$$

Lemma 4:

$$\lim_{Z \rightarrow \infty} \frac{\Phi\left(\frac{Z(\sigma)}{\sigma}\right) - \Phi\left(\frac{Z(\sigma) - \bar{\epsilon}}{\sigma}\right)}{1 - \Phi\left(\frac{Z(\sigma)}{\sigma}\right)} - \frac{\varphi\left(\frac{Z(\sigma) - \bar{\epsilon}}{\sigma}\right) - \varphi\left(\frac{Z(\sigma)}{\sigma}\right)}{\varphi\left(\frac{Z(\sigma)}{\sigma}\right)} = 0 \quad (17)$$

We can now state the main results of the paper:

Proposition 2:

$$\frac{dAR(\sigma)}{d\sigma} < \frac{\partial AR(\sigma)}{\partial \sigma} \quad \text{if and only if } Z > \hat{Z} \quad (18)$$

$$\frac{dAR(\sigma)}{d\sigma} \xrightarrow{Z \rightarrow \infty} 0 \quad \text{and} \quad \frac{\frac{dAR(\sigma)}{d\sigma}}{\frac{\partial AR(\sigma)}{\partial \sigma}} \xrightarrow{Z \rightarrow \infty} 0 \quad (19)$$

Corollary 2:

$$\frac{dAR(\sigma)}{d\sigma} \xrightarrow{C \rightarrow 0} 0 \quad \text{and} \quad \frac{\frac{dAR(\sigma)}{d\sigma}}{\frac{\partial AR(\sigma)}{\partial \sigma}} \xrightarrow{C \rightarrow 0} 0 \quad (20)$$

The first result of Proposition 2 says that the gains from improved informativeness (i.e., from a *lower* σ) are larger when assessed with the partial derivative than with the total derivative if and only if Z is larger than \hat{Z} , which is by construction the least informative performance (the likelihood ratio is equal to zero at $Z = \hat{Z}$). The second result of proposition 2 implies that, as the exercise price Z approaches infinity, improving informativeness does not decrease the agency rent. The third result implies that, as Z approaches infinity, the gains from improved informativeness are infinitely larger when assessed with the partial derivative than with the total derivative.

Because of the assumption that $\bar{\sigma}^2 - \underline{\sigma}^2$ is arbitrarily small, we can use a first-order approximation to determine whether it is optimal for a firm to filter the shock $\tilde{\epsilon}$ out of the

measure of performance $\tilde{\pi}$. It is optimal to do so if and only if

$$\frac{dAR(\sigma)}{d\sigma}(\bar{\sigma} - \underline{\sigma}) > \kappa \quad (21)$$

Corollary 2 then enables us to conclude that, for any $\kappa > 0$, it is suboptimal to set $\delta = 1$ and filter out the shock $\tilde{\epsilon}$ as $C \rightarrow 0$ (i.e., the cost of effort approaches zero). In this case, if it is costly to put in place a mechanism that enhances informativeness, then it is not optimal to do so, no matter how small the cost κ . Thus, for any $\kappa > 0$, the model predicts that pay for luck is optimal if $C \rightarrow 0$.

The reason behind this counterintuitive result is the following. A change in the variance of the performance measure has two effects on the agency rent. First, the well-known “volatility effect” is the change in the expected payoff of a call option following a change in the variance of the performance measure. Second, the “exercise price effect” is the change in the expected payoff of a call option due to the change in the exercise price $Z(\sigma)$ needed for the incentive constraint to still be satisfied as an equality following a change in σ . Interestingly, the exercise price effect can either reinforce or countervail the volatility effect. For high-powered incentives ($Z < \hat{Z}$), then $\frac{dZ(\sigma)}{d\sigma} < 0$ (see Lemma 2), so that the volatility effect and the exercise price effect both lead to a lower agency rent as σ is reduced. In particular, this is the case with a linear contract, since a linear contract is obtained in the limit, as $Z(\sigma)$ approaches $-\infty$. For low-powered incentives, however ($Z > \hat{Z}$), then $\frac{dZ(\sigma)}{d\sigma} > 0$, so that the exercise price effect partly offsets the volatility effect. In this case, reducing σ decreases the incentive-compatible exercise price $Z(\sigma)$, which increases the expected payoff of a call option contract, *ceteris paribus*, which results in a higher agency rent. Proposition 2 says that, as incentives become sufficiently low-powered ($C \rightarrow 0$, so that $Z \rightarrow \infty$), the exercise price effect fully offsets the volatility effect. In this limit case, the variance of the performance measure is irrelevant for the agency rent. This reasoning makes clear that we must consider the total effect of a change in the precision of the performance measure on the optimal contract, not just a partial effect

holding the exercise price constant.

3.2 Robustness

The existence of the exercise price effect is largely robust to the probability distribution. Indeed, in the model, as in Innes (1990), at the optimum either the limited liability constraint or the upper bound on the sensitivity of pay to performance binds (the only required distributional assumptions for this result to hold are the monotone likelihood ratio property and a support of the performance measure unbounded from above). The only contract parameter which remains to be determined is the exercise price, which adjusts to satisfy the incentive constraint as an equality – regardless of the probability distribution. With the normal distribution, the exercise price decreases following a marginal reduction in σ if and only if the likelihood ratio evaluated at $Z(\sigma)$ is positive (Lemma 2).

An important caveat is that the sign of the exercise price effect might depend on the shape of the probability distribution. This is because an increase in the variance of the performance measure reduces the density around the center of the distribution, but increases it at the tails. For a given probability distribution, the relative importance of the tails is therefore important in determining whether a change in the variance increases or decreases the agency rent. The normal distribution is known for its thin tails (or low kurtosis). In this case, the sign of the exercise price effect is given by Lemma 2 – notice that it does not depend on the variance of the distribution. As a robustness test, we compute numerically the sign of the exercise price effect for a range of parameter values with the logistic distribution, whose shape roughly resembles the normal's, but which has a markedly higher kurtosis (like the normal, this distribution also has a scale parameter which determines its variance independently of its mean; for convenience we will also refer to this scale parameter as σ). In this case, we find the same results as in Lemma 2: $\frac{dZ}{d\sigma} < 0$ for $Z < \hat{Z}$ and $\frac{dZ}{d\sigma} > 0$ for $Z > \hat{Z}$.⁹

⁹To determine the sign of $\frac{dZ}{d\sigma} > 0$, it is sufficient to determine the sign of the numerator in (32), because the denominator is negative for any probability distribution. We cannot obtain closed form solutions with the logistic distribution.

In the next three paragraphs, we provide some intuition as to why the exercise price effect might to some extent be driven by the shape of the probability distribution, and notably its kurtosis. We argue that the sign of the exercise price effect when Z is low is in principle ambiguous, but that the sign of the exercise price effect when Z is high is largely robust to the probability distribution. To start with, note that the sensitivity of the exercise price to σ is driven by the following mechanism: the exercise price adjusts to a change in σ so that the incentive constraint is still satisfied as an equality. That is, if the difference between expected pay with $e = \bar{e}$ and $e = 0$ increases following a change in σ (which raises incentives), then the exercise price increases (which lowers incentives), and conversely.

First, suppose that Z is “low”. An increase in σ removes probability mass from the center of any distribution – whether the one corresponding to $e = \bar{e}$ or to $e = 0$. In addition, the effect on $E[W(\tilde{\pi})|e]$ around the center of each distribution is stronger for $e = \bar{e}$ than for $e = 0$, because the former ascribes higher probability mass to high payments relative to low payments than the latter (this is implied by the monotone likelihood ratio property and the fact that the compensation contract is increasing in performance¹⁰). That is, if for the sake of the argument we only consider the “center” of the distribution, an increase in σ reduces $E[W(\tilde{\pi})|e = \bar{e}]$ more than $E[W(\tilde{\pi})|e = 0]$. This effect alone suggests an inverse relationship between σ and Z (i.e., $\frac{dZ}{d\sigma} < 0$) for sufficiently low Z . Furthermore, this effect is all the more important that probability mass is concentrated around the center of the distribution (i.e., the distribution has thin tails, as is the case with the normal). An opposing effect is that $E[W(\tilde{\pi})|e]$ increases at the tails of the distribution following an increase in σ . Moreover, again because of the monotone likelihood ratio property and the increasing contract, the increase in $E[W(\tilde{\pi})|e = \bar{e}]$ is larger than the increase in $E[W(\tilde{\pi})|e = 0]$. Once again, this is because the density associated with $e = \bar{e}$ ascribes higher probability mass to high payments relative to low payments than the density associated with $e = 0$, both at the left and the right tails of the distribution. This effect alone suggests a positive relationship between σ and Z (i.e.,

¹⁰Notice that the monotone likelihood ratio property is sufficient here: if it holds, then the compensation contract is increasing.

$\frac{dZ}{d\sigma} > 0$) for sufficiently low Z . Nevertheless, we have found that the first effect still dominates with the logistic distribution, in spite of its thicker tails.

Second, for a higher Z , the first effect mentioned in the paragraph above, which operates at the center of the distribution, is associated with low payments (because the contract is increasing), so that it is outweighed by the second effect mentioned above (the tails effect).¹¹ This suggests that the result that $\frac{dZ}{d\sigma} > 0$ for sufficiently high Z is robust to the probability distribution. Indeed, it can be shown that if effort results in a rightward shift in the probability distribution of performances, and if the density is convex above a certain threshold (which we denote by \underline{Z}), then $\frac{dZ}{d\sigma} > 0$ for any $Z \geq \max\{\bar{e}, \underline{Z}\}$. This argument suggests that our result that the overall gains (with the volatility effect and the exercise price effect) from improved informativeness are lower than the direct gains (with the volatility effect only) for sufficiently low-powered incentives is robust to the probability distribution, as long as certain conditions are satisfied.¹²

3.3 Numerical simulations

It is also instructive to assess numerically the gains from improved informativeness for a range of values of Z , and to compare these gains to “naïve” assessments which do not take into account the exercise price effect. For computational reasons and consistency, we only consider the normal distribution in what follows.

In Figures 1 and 2, we report both the value of $\frac{d}{d\sigma}AR(\sigma)$ (“total sensitivity”), as calculated in (39) in the Appendix, and the value of $\frac{\partial}{\partial\sigma}AR(\sigma)$ (“partial sensitivity”), as calculated in (35), for a range of values of Z (the exercise price Z is a measure of the inverse of the power of incentives, see Corollary 1) and different parameter values: Figure 1 uses $\sigma = 1$, and Figure 2 uses $\sigma = 3$. That is, Figures 1 and 2 display the gains from improved informativeness for

¹¹For a sufficiently high Z , the first effect is irrelevant, i.e., the density increases following an increase in σ at every performance associated with a strictly positive payment.

¹²The condition that effort results in a rightward shift in the probability distribution means that effort increases expected profits but does not affect higher-order moments of the distribution of profits; the convexity condition is satisfied by any continuous distribution with unbounded support.

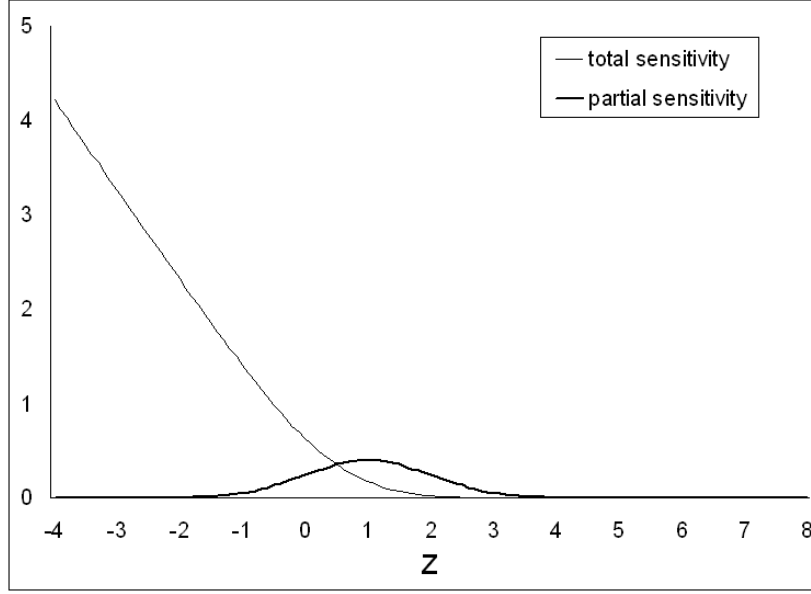


Figure 1: Sensitivity of the agency rent to σ for a range of values of Z , for $b = \bar{e} = 1$ and $\sigma = 1$. The total (respectively partial) sensitivity is the total (resp. partial) derivative of the agency rent with respect to σ .

different values of Z and σ . Note that with $\bar{e} = 1$, $\hat{Z} = 0.5$. The “total sensitivity” represents the overall gains from improved informativeness, if both the volatility effect and the exercise price effect are taken into account. It follows from (21) that the total sensitivity function on Figures 1 and 2 can also be viewed as the maximum value of $\frac{\kappa}{\bar{\sigma} - \underline{\sigma}}$ such that it is optimal to filter out the shock $\tilde{\epsilon}$. The “partial sensitivity” represents the partial gains from improved informativeness, if only the volatility effect is taken into account – so that the induced change in the form of the optimal compensation contract is omitted.

Since both the partial sensitivity and the total sensitivity are linear in b (as can be seen in (39)), the shapes of the functions displayed on Figures 1 and 2 do not depend on b . Numerical simulations also show that neither the gains nor the shapes of these functions depend strongly on \bar{e} .

The gains with the “partial sensitivity” measure tend to zero as the exercise price approaches either $-\infty$ or ∞ . This can be explained as follows. As $Z \rightarrow -\infty$, the contract

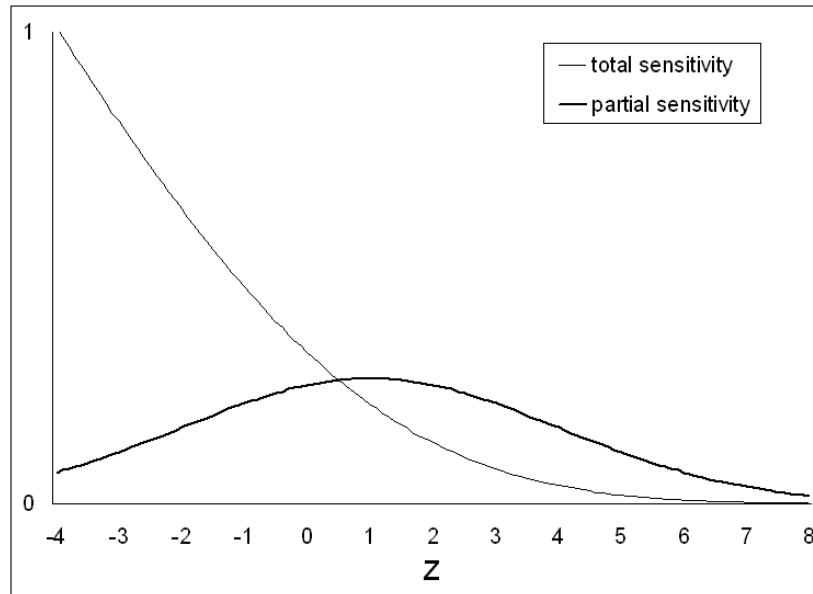


Figure 2: Sensitivity of the agency rent to σ for a range of values of Z , for $b = \bar{e} = 1$ and $\sigma = 3$. The total (respectively partial) sensitivity is the total (resp. partial) derivative of the agency rent with respect to σ .

becomes approximately linear. Besides, with a symmetrically distributed performance measure, we know that the payoff of a linear contract is independent from the volatility of this performance measure. Likewise, as $Z \rightarrow \infty$, compensation becomes approximately constant, so that it is not sensitive to changes in the volatility of the performance measure. On the contrary, the gains are maximum for $Z = \bar{e}$, i.e., when the exercise price is equal to the expected performance. In this case, holding the form of the compensation contract constant, an increase in the volatility of the performance increases compensation conditional on a “good” performance ($\pi > \bar{e}$), but it does not decrease compensation conditional on a “bad” performance ($\pi < \bar{e}$). In short, the relationship between the power of incentives and the gains from improved informativeness as assessed with the “partial sensitivity” measure are nonmonotonic, and these gains are maximum for at-the-money options ($Z = \bar{e}$).¹³

¹³Options are “at-the-money” when the exercise price is equal to the stock price at the time when they are granted. Suppose that a stock which entitles his holder to the firm profits at $t = 1$ is traded at $t = -1$. With no time discounting and risk neutral investors with rational expectations, the stock price at $t = 0$ would be equal to expected profits, i.e., $\bar{e} = 1$, in equilibrium. With this minor and straightforward extension of the

As in Proposition 2, we find that only taking into account the partial sensitivity (i.e., omitting the exercise price effect) leads to an underestimation of the total gains from improved informativeness for $Z < \hat{Z}$, but it leads to an overestimation of the total gains from improved informativeness for $Z > \hat{Z}$. As already discussed, this is because in the former case the exercise price effect reinforces the impact of the volatility effect, and in the latter case the exercise price effect partly offsets the volatility effect.

With the total sensitivity measure, the gains from increased informativeness are especially high for a low Z , i.e., for high-powered incentives. On the contrary, with the partial sensitivity measure, the gains from increased informativeness are maximal for $Z = \bar{e}$, i.e., for an intermediate level of incentives – at $Z = \bar{e}$, the exercise price is equal to the expected value $E[\tilde{\pi}]$ of the performance measure in equilibrium (this immediately follows from (1)). Furthermore, the lower σ , the greater the difference in gains under these two sensitivity measures at $Z = \bar{e}$. For example, with $\sigma = 3$, the gains are only 20.6% lower with the total sensitivity measure than with the partial sensitivity measure; for $\sigma = 0.5$, the gains are 10.6 times lower with the total sensitivity measure than with the partial sensitivity measure.

Although not clearly perceptible on the graphs, the speed of convergence of the “total sensitivity” function toward zero is remarkable. For $\sigma = 1$, at $Z = 5$ the gains from a marginal change in σ are more than two millions times smaller than at $Z = 0$. For $\sigma = 3$, however, the gains are “only” 16.9 times smaller at $Z = 5$ than at $Z = 0$. In this case, a two-millions-to-one ratio only obtains for $Z \approx 14.56$.

Moreover, and crucially, the partial sensitivity measure vastly overestimates the total gains from improved informativeness for sufficiently large Z . For example, for $\sigma = 1$ and $Z = 2$ (which is only one standard deviation away from the expected performance of $\bar{e} = 1$), the gains from a marginal change in σ are 10.7 times larger with the partial sensitivity measure than with the total sensitivity measure. For a given σ , this ratio grows steadily with Z : for $\sigma = 1$ and $Z = 5$, this ratio is 483.8. This finding is important, because it shows that even

model, we can refer to options as being “at”, “in”, or “out” of the money.

for a range of *non-extreme* parameter values, gains from improved informativeness can be much lower if the exercise price effect is taken into account. This being said, the discrepancy seems to narrow as σ is increased: for $\sigma = 3$ and $Z = 4$, the same ratio is 4.1; for $\sigma = 3$ and $Z = 14.56$, it is 60.6.

It is also instructive to compute the gains from improved informativeness with both sensitivity measures for a range of exercise prices as a *percentage* of the expected cost of compensation given an exercise price. This is illustrated in Figure 3, for $\sigma = 1$ and a 1% reduction in σ .¹⁴ Obviously, for any Z , the ordering of the partial and total sensitivity measures remains unchanged. However, when expressed as a fraction of the expected cost of compensation, the difference between gains for a high Z and gains for a low Z increases relative to the case where gains are measured in absolute terms, since the expected cost of compensation for a high Z is much lower than for a low Z . The reason why gains as estimated with the partial sensitivity measure are increasing in the exercise price is clear: the sensitivity of the price of a call option to volatility, when expressed as a fraction of the call price, increases with the exercise price. We can see on Figure 3 that the percentage gains from improved informativeness, as estimated with the partial sensitivity measure, can be extremely high for out-of-the-money options ($Z > 1$), but extremely low for in-the-money options ($Z < 1$). In comparison, the percentage gains estimated with the total sensitivity measure can be extremely low for out-of-the-money options (for $Z = 5$, the gains associated with a 1% reduction in σ are equal to 0.04% of the expected cost of compensation), but they are moderate for in-the-money options. For at-the-money options, gains are equal to 1.00% of the expected cost of compensation with the partial sensitivity measure, and to 0.42% with the total sensitivity measure. It should be noted that these figures are obtained in a highly stylized model and should not be overinterpreted.

In summary, these numerical results establish that there can be a wide discrepancy between

¹⁴Gains are estimated at the margin, so that the reduction in σ cannot be too high for the approximation with first-order terms to remain acceptable. This being said, gains for a 10% reduction in σ , say, can be estimated by extrapolation – they will be approximately equal to ten times the gains from a 1% reduction in σ .

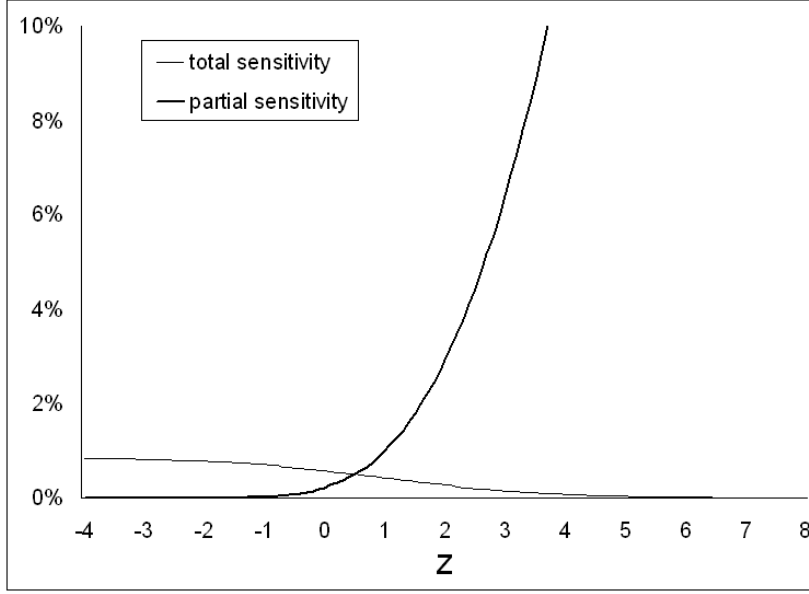


Figure 3: Change in the agency rent for a 1% change in σ , expressed as a fraction of the expected cost of compensation in equilibrium, for $b = \bar{e} = 1$ and $\sigma = 1$. The total (respectively partial) sensitivity is the total (resp. partial) derivative of the agency rent with respect to σ .

the partial and total sensitivity measures of gains from improved informativeness, and that the partial sensitivity measure will overestimate the total gains in some cases, and underestimate them in other cases. More precisely, when incentives are sufficiently low-powered, not taking into account the exercise price effect could lead to considerably overestimate the value of a reduction in the noise of the performance measure. To the extent that there exists even small costs to indexing compensation contracts, this result can explain the lack of indexing and the presence of pay-for-luck in compensation contracts. This result also emphasizes that it is important to take into consideration changes in the form of the optimal compensation contract induced by a change in the noise of the performance measure.

3.4 Predictions

The model predicts that RPE and other benchmarking techniques which improve the informativeness of the performance measure will be more prevalent when incentives are high-powered

(low Z), *ceteris paribus*. That is, employees with higher-powered incentives such as CEOs will be less paid for luck and will be more subject to relative performance evaluation than other employees. Put differently, employees with low-powered incentives will be more paid for luck. Besides, this prediction holds whether filtering out shocks to performance for any given employee is a fixed cost (Figure 1) or is proportional to the expected cost of compensation (Figure 3). This prediction is consistent with the results of Oyer and Schaefer (2005), who cannot explain broad stock-options plans with a standard “informativeness agency model.” Moreover, for a given cost of improving informativeness, the model predicts that the need to provide stronger incentives could result in a higher informativeness of the performance measure.

Interestingly, an analysis which focuses exclusively on the partial sensitivity measure would reach the opposite conclusion, i.e., that indexing and benchmarking are most valuable when incentives are low-powered (high Z), for example in the form of an unlikely reward for very high performances.

4 Conclusion

In this paper, we show that changes in the form of the optimal contract induced by a change in the noise of the performance measure can significantly affect the value of improved informativeness. To this end, we use a standard principal-agent model in the spirit of Innes (1990). In this setting, a compensation contract taking the form of a call option on the performance measure is optimal for any given variance of the performance measure. We show both theoretically and numerically that *not* taking into account changes in the form of the optimal contract would lead to largely underestimate the gains from improved informativeness in some cases, and to largely overestimate them in other cases. Thus, the model can contribute to explain two stylized facts of executive compensation: the use of the stock price as a performance measure, and the use of options on the stock price. More generally, our

results provide a novel potential explanation for violations of the informativeness principle reported in the literature.

It may be possible to generalize the main result of the paper to a broader class of probability distributions than the normal, or to identify other sets of modeling assumptions under which similar results hold – and these are natural avenues for future research. Specific results can naturally be expected to depend on the setting used. This however does not affect the key messages of this paper that it is essential to take into account changes in the form of the optimal contract to assess the value of increased informativeness, especially in the presence of constraints on contracting, and that neglecting this effect would in some instances lead to largely overestimate it.

5 Bibliography

Aggarwal R.K., Samwick A.A. 1999. Executive compensation, relative performance evaluation, and strategic competition: theory and evidence. *Journal of Finance*, 54 (6), 1999-2043.

Axelson, Ulf, and Sandeep Baliga, 2009, Liquidity and manipulation of executive compensation schemes, *Review of Financial Studies*, 22 (10), 3907-3939.

Bertrand, Marianne, and Sendhil Mullainathan, 2001, Are CEOs rewarded for luck? The ones without principals are, *Quarterly Journal of Economics*, 116 (3), 901-932.

Bolton, Patrick and Mathias Dewatripont, 2005, *Contract Theory* (MIT Press).

Core, J.E., Guay, W.R., 2001, Stock option plans for non-executive employees, *Journal of Financial Economics*, 60, 253-267.

Danthine, Jean-Pierre, and John B. Donaldson, 2008, Executive compensation: a general equilibrium perspective, working paper.

Dewatripont, Mathias, Patrick Legros, and Steven A. Matthews, 2003, Moral hazard and capital structure dynamics, *Journal of the European Economic Association*, 1 (4), 890-930.

Dittmann, Ingolf, and Ernst Maug, 2007, Lower salaries and no options? On the optimal

structure of executive pay, *Journal of Finance* 62 (1), 303-343.

Dittmann, Ingolf, Ernst Maug, and Oliver Spalt, 2011, Indexing executive compensation contracts, Working paper.

Edmans, A., Gabaix, X., 2009, Is CEO Pay Really Inefficient? A Survey of New Optimal Contracting Theories, *European Financial Management*, 15 (3), 486-496.

Frydman, Carola, and Dirk Jenter, 2010, CEO compensation, *Annual Review of Financial Economics*, 2 (1), 75-102.

Garvery, G.T., Milbourn, T.T., 2006, Asymmetric benchmarking in compensation: Executives are rewarded for good luck but not penalized for bad, *Journal of Financial Economics*, 82 (1), 197-225.

Gjesdal, Froystein, 1982, Information and incentives: the agency information problem, *Review of Economic Studies*, 49 (3), 373-390.

Gopalan, R., Milbourn, T. and Song, F., Strategic flexibility and the optimality of pay for luck, *Review of Financial Studies*, 23 (5): 2060-2098.

Grossman, Sanford J., and Oliver D. Hart, 1983, An analysis of the principal-agent problem, *Econometrica* 51 (1), 7-45.

Hall, Brian J., Kevin J. Murphy, 2002, Stock options for undiversified executives, *Journal of Accounting and Economics* 33, 3-42.

Holmstrom, Bengt, 1979, Moral hazard and observability, *Bell Journal of Economics* 10 (1), 74-91.

Hemmer, Thomas, Oliver Kim, and Robert E. Verrecchia, 2000, Introducing convexity into optimal compensation contracts, *Journal of Accounting and Economics* 28, 307-327.

Innes, Robert D., 1990, Limited liability and incentive contracting with ex-ante action choices, *Journal of Economic Theory* 52, 45-67.

Kim, Son Ku, 1995, Efficiency of an information system in an agency model, *Econometrica*, 63 (1), 89-102.

Laffont, Jean-Jacques and David Martimort, 2002, *The Theory of Incentives* (Princeton

University Press).

Murphy, K.J. “Executive compensation, in *Handbook of Labor Economics*, O. Ashenfelter and D. Card, eds. North Holland (1999).

Oyer, Paul, 2004, Why do firms use incentives that have no incentive effects?, *Journal of Finance*, 59 (4), 1619-1649.

Oyer, Paul, and Scott Schaefer, 2005, Why do firms give stock options to all employees?: An empirical examination of alternative theories, *Journal of Financial Economics*, 76, 99-133.

Raith, Michael, 2008, Specific knowledge and performance measurement, *Rand Journal of Economics*, 39 (4), 1059-1079.

Salanié, B., 1997, *The Economics of Contracts* (MIT Press).

Shavell, Steven, 1979, Risk sharing and incentives in the principal and agent relationship, *Bell Journal of Economics*, 10 (1), 55-73.

Tirole, Jean, 2006. *The Theory of Corporate Finance* (Princeton University Press).

6 Appendix

Proof of Proposition 1:

We are going to rely on the fact that $W(\pi)$ is continuous, which may be justified as follows. On the one hand, the likelihood ratio is increasing in π for any π , so that the optimal contract is nondecreasing in π for any π , which rules out downward discontinuities in $W(\pi)$. On the other hand, the constraint that $W'(\pi) \leq 1$ rules out upward discontinuities.

Denoting the Lagrange multipliers associated to the four constraints (9), (10), (11), and (12) for a given π by μ , $\lambda(\pi)$, $\gamma(\pi)$, and $\eta(\pi)$ respectively, the Lagrangian writes as

$$\begin{aligned} L = & \int_{-\infty}^{\infty} W(\pi)\psi(\pi|e=0)d\pi - \mu \left[\int_{-\infty}^{\infty} W(\pi)\psi(\pi|e=\bar{e})d\pi - C - \int_{-\infty}^{\infty} W(\pi)\psi(\pi|e=0)d\pi \right] \\ & - \int_{-\infty}^{\infty} \lambda(\pi)W(\pi)d\pi - \int_{-\infty}^{\infty} \gamma(\pi)(\pi - W(\pi))d\pi - \int_{-\infty}^{\infty} \eta(\pi)(b - W'(\pi))d\pi \end{aligned}$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \left((1 + \mu)W(\pi) + \mu C - \mu \frac{\psi(\pi|e = \bar{e})}{\psi(\pi|e = 0)} W(\pi) \right) \psi(\pi|e = 0) d\pi \\
&\quad - \int_{-\infty}^{\infty} \left(\lambda(\pi)W(\pi) + \gamma(\pi)(\pi - W(\pi)) - \eta(\pi)(b - W'(\pi)) \right) d\pi
\end{aligned} \tag{22}$$

We use the facts that the optimal contract $W(\pi)$ is continuous and $\lim_{\pi \rightarrow -\infty} W(\pi) = 0$ ¹⁵ to write

$$W(\pi) = \int_{-\infty}^{\pi} W'(x) dx = \int_{-\infty}^{\pi-d\pi} W'(x) dx + W'(\pi) d\pi \tag{23}$$

It follows that

$$\frac{dW'(\pi)}{dW(\pi)} = \left(\frac{dW(\pi)}{dW'(\pi)} \right)^{-1} = \left(\frac{d(\int_{-\infty}^{\pi-d\pi} W'(x) dx + W'(\pi) d\pi)}{dW'(\pi)} \right)^{-1} = \frac{1}{d\pi} \tag{24}$$

The first-order necessary condition with respect to $W(\pi)$ to the constrained optimization problem is

$$\left((1 + \mu) + -\mu \frac{\psi(\pi|e = \bar{e})}{\psi(\pi|e = 0)} \right) \psi(\pi|e = 0) - \lambda(\pi) + \gamma(\pi) + \eta(\pi) \frac{1}{d\pi} = 0 \tag{25}$$

We define

$$\phi(\pi) \equiv \left((1 + \mu) + -\mu \frac{\psi(\pi|e = \bar{e})}{\psi(\pi|e = 0)} \right) \psi(\pi|e = 0) \tag{26}$$

If $\phi(\pi) > 0$ for a given π , then (25) imposes that

$$-\lambda(\pi) + \gamma(\pi) + \eta(\pi) \frac{1}{d\pi} < 0$$

Given the nonnegativity constraints on the Lagrange multipliers, a necessary condition for this inequality to hold is $\lambda(\pi) > 0$. But $\lambda(\pi) > 0$ if and only if $W(\pi) = 0$.

¹⁵Because the likelihood ratio of the normal distribution is increasing, we know that the optimal contract is nondecreasing in π . Therefore, $W(\pi) \geq \lim_{x \rightarrow -\infty} W(x)$, for any π . Consider any contract $W(\pi)$ characterized by $\lim_{x \rightarrow -\infty} W(x) = \varepsilon > 0$. Then define $W^*(\pi) \equiv W(\pi) - \varepsilon$ for all π . The contract $W^*(\pi)$ does not violate any constraint (in particular, it satisfies the incentive constraint), and it is characterized by a lower agency rent than $W(\pi)$. This shows that the contract $W(\pi)$ is not optimal.

If $\phi(\pi) < 0$ for a given π , then (25) imposes that

$$-\lambda(\pi) + \gamma(\pi) + \eta(\pi)\frac{1}{d\pi} > 0 \quad (27)$$

Given the nonnegativity constraints on the Lagrange multipliers, either $\gamma(\pi) > 0$ and $\eta(\pi) > 0$, or $\gamma(\pi) > 0$ and $\eta(\pi) = 0$, or $\gamma(\pi) = 0$ and $\eta(\pi) > 0$.

Since μ is a constant and the likelihood ratio is monotonically increasing in π , we know from (26) that $\phi(\pi) = 0$ on a set of probability zero, and that the optimal contract (which is incentive-compatible) exists if there exists a finite and positive Z such that

$$\phi(\pi) > 0 \quad \Leftrightarrow \quad \frac{\psi(\pi|e = \bar{e})}{\psi(\pi|e = 0)} < \frac{1 + \mu}{\mu} \quad \Leftrightarrow \quad \pi < Z \quad (28)$$

$$\phi(\pi) < 0 \quad \Leftrightarrow \quad \frac{\psi(\pi|e = \bar{e})}{\psi(\pi|e = 0)} > \frac{1 + \mu}{\mu} \quad \Leftrightarrow \quad \pi > Z \quad (29)$$

We also know that $W(\pi) = 0$ for $\pi < Z$. We still have to determine the form of the optimal contract for $\pi > Z$. Given the nonnegativity constraints on the Lagrange multipliers discussed above, we know that, for $\pi > Z$, either (11) is binding, or (12) is binding, or both are binding. We now show by contradiction that only (12) is binding for $\pi > Z$.

Suppose that, for any given $\pi \in [Z, m)$ (with any given $m > Z$), $W(\pi) = b\pi$. Then we must have $m = \infty$. Otherwise, for some $\pi > m$, we would have $W(\pi) < b\pi$ and $W'(\pi) < b$, which is impossible since it would imply that $\gamma(\pi) = 0$ and $\eta(\pi) = 0$, which would violate (27). However, to have $W(\pi) = b\pi$ for any $\pi \in [n, \infty)$, where $n \geq Z$, subject to the constraint $W'(\pi) \leq b$, we need to have $Z = 0$. However, a contract with $Z = 0$ and $W(\pi) = b\pi$ for any $\pi > Z$ does not bind the incentive constraint, because of (7), which is suboptimal. Therefore, the optimal contract cannot be characterized by $W(\pi) = b\pi$, for any $\pi > Z$. This implies that $\gamma(\pi) = 0$ for any $\pi > Z$, which in turn implies that $\eta(\pi) > 0$ for any $\pi > Z$, so that $W'(\pi) = b$ for any $\pi > Z$.

The optimal contract is therefore characterized by $W(\pi) = 0$ for all $\pi < Z$, and $W'(\pi) = b$

for all $\pi > Z$. Because of (10) and (11), we know that Z cannot be negative.

Proof of Lemma 1:

First, for a contract characterized by $W(\pi) = 0$ for all $\pi < Z$, and $W'(\pi) = b$ for all $\pi > Z$, the incentive constraint with a given Z is

$$\int_Z^\infty b(\pi - Z)\psi(\pi|e = \bar{e})d\pi - \int_Z^\infty b(\pi - Z)\psi(\pi|e = 0)d\pi = C \quad (30)$$

The first-derivative of the left-hand-side of (59) with respect to Z is equal to

$$\begin{aligned} & -b \int_Z^\infty \psi(\pi|e = \bar{e})d\pi + b \int_Z^\infty \psi(\pi|e = 0)d\pi \\ & = -bPr(\pi > Z|e = \bar{e}) + bPr(\pi > Z|e = 0) \end{aligned}$$

which is strictly negative for any Z , because of first-order stochastic dominance, which is implied by a monotone likelihood ratio.

Second, the left-hand side of (59) tends to zero as Z approaches ∞ .

Third, (7) guarantees that the left-hand-side of (59) is strictly larger than C for $Z = 0$.

Therefore, there exists one and only one positive and finite Z such that the left-hand-side of (59) equals C , i.e., such that the incentive-constraint with the optimal contract is satisfied as an equality.

Proof of Corollary 1: On the one hand, $Z(\sigma)$ satisfies the incentive constraint in (59) as an equality, by definition. On the other hand, the left-hand-side of the incentive constraint in (59) is strictly decreasing in Z (see Lemma 1). Since the right-hand-side of (59) is equal to C , this implies that $Z(\sigma)$ is strictly decreasing in C .

Proof of Lemma 2 and Lemma 3:

We know that $E[\tilde{\pi}] = 0$ for $e = 0$, $E[\tilde{\pi}] = \bar{e}$ for $e = \bar{e}$, and $var[\tilde{\pi}] = \sigma^2$ for any e . When $W(\pi)$ is as in (14), the agency rent is

$$AR(\sigma) = E[W(\tilde{\pi})|e = 0] = \int_{Z(\sigma)}^{\infty} b(x - Z(\sigma)) \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} dx$$

Taking the total derivative of the agency rent with respect to σ ,

$$\frac{d}{d\sigma} AR(\sigma) = \frac{d}{d\sigma} E[W(\tilde{\pi})|e = 0] = \frac{\partial}{\partial \sigma} E[W(\tilde{\pi})|e = 0] + \frac{\partial}{\partial Z} \{E[W(\tilde{\pi})|e = 0]\} \frac{dZ(\sigma)}{d\sigma} \quad (31)$$

For a given σ , marginally increasing σ results in a higher agency rent in equilibrium if and only if this expression is positive.

For any change in σ , $Z(\sigma)$ adjusts so that the incentive constraint in (59) is still satisfied as an equality, by definition of $Z(\sigma)$. The implicit function theorem yields

$$\frac{dZ(\sigma)}{d\sigma} = - \frac{\frac{\partial \{E[W(\tilde{\pi})|e=\bar{e}] - E[W(\tilde{\pi})|e=0]\}}{\partial \sigma}}{\frac{\partial \{E[W(\tilde{\pi})|e=\bar{e}] - E[W(\tilde{\pi})|e=0]\}}{\partial Z}} \quad (32)$$

Substituting in (31),

$$\frac{d}{d\sigma} AR(\sigma) = \frac{d}{d\sigma} E[W(\tilde{\pi})|e = 0] = \frac{\partial}{\partial \sigma} E[W(\tilde{\pi})|e = 0] - \frac{\partial}{\partial Z} \{E[W(\tilde{\pi})|e = 0]\} \frac{\frac{\partial \{E[W(\tilde{\pi})|e=\bar{e}] - E[W(\tilde{\pi})|e=0]\}}{\partial \sigma}}{\frac{\partial \{E[W(\tilde{\pi})|e=\bar{e}] - E[W(\tilde{\pi})|e=0]\}}{\partial Z}} \quad (33)$$

First,

$$\begin{aligned} \frac{\partial}{\partial \sigma} E[W(\tilde{\pi})|e] &= \frac{\partial}{\partial \sigma} \int_{Z(\sigma)}^{\infty} b(x - Z(\sigma)) \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x - e)^2}{2\sigma^2}\right\} dx \\ &= \frac{\partial}{\partial \sigma} \int_{Z(\sigma)-e}^{\infty} b(x + e - Z(\sigma)) \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} dx \\ &= \frac{\partial}{\partial \sigma} \int_{Z(\sigma)-e}^{\infty} bx \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} dx - b(Z(\sigma) - e) \frac{\partial}{\partial \sigma} \int_{Z(\sigma)-e}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} dx \\ &= b \frac{\partial}{\partial \sigma} \left\{ \frac{\sigma}{\sqrt{2\pi}} \exp\left\{-\frac{(Z(\sigma) - e)^2}{2\sigma^2}\right\} \right\} - b(Z(\sigma) - e) \frac{\partial}{\partial \sigma} \int_{\frac{Z(\sigma)-e}{\sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du \end{aligned} \quad (34)$$

$$\begin{aligned}
&= \frac{b}{\sqrt{2\pi}} \exp \left\{ -\frac{(Z(\sigma) - e)^2}{2\sigma^2} \right\} \left[1 + \frac{(Z(\sigma) - e)^2}{\sigma^2} \right] - \frac{b}{\sqrt{2\pi}} \frac{(Z(\sigma) - e)^2}{\sigma^2} \exp \left\{ -\frac{(Z(\sigma) - e)^2}{2\sigma^2} \right\} \\
&= \frac{b}{\sqrt{2\pi}} \exp \left\{ -\frac{(Z(\sigma) - e)^2}{2\sigma^2} \right\} = b\varphi\left(\frac{Z(\sigma) - e}{\sigma}\right)
\end{aligned} \tag{35}$$

which is positive.

Second,

$$\begin{aligned}
\frac{\partial}{\partial Z} E[W(\tilde{\pi})|e] &= \frac{\partial}{\partial Z} \int_{Z(\sigma)}^{\infty} b(x - Z(\sigma)) \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - e)^2}{2\sigma^2} \right\} dx \\
&= -b \int_{Z(\sigma)}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - e)^2}{2\sigma^2} \right\} dx = -b \left(1 - \Phi \left(\frac{Z(\sigma) - e}{\sigma} \right) \right)
\end{aligned} \tag{36}$$

Using (35) and (36), (32) rewrites as

$$\frac{dZ(\sigma)}{d\sigma} = - \frac{\left[\varphi\left(\frac{Z(\sigma) - \bar{e}}{\sigma}\right) - \varphi\left(\frac{Z(\sigma)}{\sigma}\right) \right]}{-(1 - \Phi\left(\frac{Z(\sigma) - \bar{e}}{\sigma}\right)) + (1 - \Phi\left(\frac{Z(\sigma)}{\sigma}\right))} \tag{37}$$

The monotone likelihood ratio property implies first-order stochastic dominance, therefore

$$\Phi\left(\frac{Z(\sigma)}{\sigma}\right) > \Phi\left(\frac{Z(\sigma) - \bar{e}}{\sigma}\right)$$

so that (37) is positive if and only if

$$\varphi\left(\frac{Z(\sigma) - \bar{e}}{\sigma}\right) > \varphi\left(\frac{Z(\sigma)}{\sigma}\right) \tag{38}$$

But since

$$\varphi\left(\frac{\hat{Z} - \bar{e}}{\sigma}\right) \equiv \varphi\left(\frac{\hat{Z}}{\sigma}\right)$$

by definition of \hat{Z} , and since two p.d.f. of normal variables with the same variance intersect only once, we know that (38) is satisfied, and $\frac{dZ(\sigma)}{d\sigma} > 0$, if and only if $Z(\sigma) > \hat{Z}$. Lemma 2 is proven.

Using (35) and (36), we can also rewrite (33) as

$$\begin{aligned}
\frac{d}{d\sigma}AR(\sigma) &= b\varphi\left(\frac{Z(\sigma)}{\sigma}\right) + b\left(1 - \Phi\left(\frac{Z(\sigma)}{\sigma}\right)\right) \frac{\left[\varphi\left(\frac{Z(\sigma)-\bar{e}}{\sigma}\right) - \varphi\left(\frac{Z(\sigma)}{\sigma}\right)\right]}{-(1 - \Phi\left(\frac{Z(\sigma)-\bar{e}}{\sigma}\right)) + (1 - \Phi\left(\frac{Z(\sigma)}{\sigma}\right))} \\
&= b\varphi\left(\frac{Z(\sigma)}{\sigma}\right) - b\frac{1 - \Phi\left(\frac{Z(\sigma)}{\sigma}\right)}{\Phi\left(\frac{Z(\sigma)}{\sigma}\right) - \Phi\left(\frac{Z(\sigma)-\bar{e}}{\sigma}\right)} \left[\varphi\left(\frac{Z(\sigma)-\bar{e}}{\sigma}\right) - \varphi\left(\frac{Z(\sigma)}{\sigma}\right)\right]
\end{aligned} \tag{39}$$

This expression is strictly positive if and only if

$$\frac{\varphi\left(\frac{Z(\sigma)-\bar{e}}{\sigma}\right) - \varphi\left(\frac{Z(\sigma)}{\sigma}\right)}{\varphi\left(\frac{Z(\sigma)}{\sigma}\right)} < \frac{\Phi\left(\frac{Z(\sigma)}{\sigma}\right) - \Phi\left(\frac{Z(\sigma)-\bar{e}}{\sigma}\right)}{1 - \Phi\left(\frac{Z(\sigma)}{\sigma}\right)}$$

The proof of Lemma 3 is complete.

Proof of Lemma 4:

Define

$$f(Z) \equiv \left(\varphi\left(\frac{Z-\bar{e}}{\sigma}\right) - \varphi\left(\frac{Z}{\sigma}\right)\right)\left(1 - \Phi\left(\frac{Z}{\sigma}\right)\right) \quad \text{and} \quad g(Z) \equiv \left(\Phi\left(\frac{Z}{\sigma}\right) - \Phi\left(\frac{Z-\bar{e}}{\sigma}\right)\right)\varphi\left(\frac{Z}{\sigma}\right) \tag{40}$$

We need to show that

$$\lim_{Z \rightarrow \infty} g(Z) - f(Z) \leq 0 \tag{41}$$

The first derivative of $f(Z)$ with respect to $\frac{Z}{\sigma}$ is

$$\left(\varphi'\left(\frac{Z-\bar{e}}{\sigma}\right) - \varphi'\left(\frac{Z}{\sigma}\right)\right)\left(1 - \Phi\left(\frac{Z}{\sigma}\right)\right) - \left(\varphi\left(\frac{Z-\bar{e}}{\sigma}\right) - \varphi\left(\frac{Z}{\sigma}\right)\right)\varphi\left(\frac{Z}{\sigma}\right) \tag{42}$$

$$= \left(-\frac{Z-\bar{e}}{\sigma}\varphi\left(\frac{Z-\bar{e}}{\sigma}\right) + \frac{Z}{\sigma}\varphi\left(\frac{Z}{\sigma}\right)\right)\left(1 - \Phi\left(\frac{Z}{\sigma}\right)\right) - \left(\varphi\left(\frac{Z-\bar{e}}{\sigma}\right) - \varphi\left(\frac{Z}{\sigma}\right)\right)\varphi\left(\frac{Z}{\sigma}\right) \tag{43}$$

The first derivative of $g(Z)$ with respect to $\frac{Z}{\sigma}$ is

$$\varphi'\left(\frac{Z}{\sigma}\right)\left(\Phi\left(\frac{Z}{\sigma}\right) - \Phi\left(\frac{Z-\bar{e}}{\sigma}\right)\right) + \varphi\left(\frac{Z}{\sigma}\right)\left(\varphi\left(\frac{Z}{\sigma}\right) - \varphi\left(\frac{Z-\bar{e}}{\sigma}\right)\right) \tag{44}$$

$$= -\frac{Z}{\sigma}\varphi\left(\frac{Z}{\sigma}\right)\left(\Phi\left(\frac{Z}{\sigma}\right) - \Phi\left(\frac{Z-\bar{e}}{\sigma}\right)\right) + \varphi\left(\frac{Z}{\sigma}\right)\left(\varphi\left(\frac{Z}{\sigma}\right) - \varphi\left(\frac{Z-\bar{e}}{\sigma}\right)\right) \quad (45)$$

Using (43) and (45), we can write

$$g'(Z) - f'(Z) = (-g(Z) + f(Z))\frac{Z}{\sigma} - \frac{\bar{e}}{\sigma}\varphi\left(\frac{Z-\bar{e}}{\sigma}\right)\left(1 - \Phi\left(\frac{Z}{\sigma}\right)\right) \quad (46)$$

We first show by contradiction that $\lim_{Z \rightarrow \infty} g(Z) - f(Z) < 0$. Because of the asymptotic properties of the normal distribution's p.d.f. and c.d.f., (46) implies that

$$\lim_{Z \rightarrow \infty} g'(Z) - f'(Z) = (-g(Z) + f(Z))\frac{Z}{\sigma} \quad (47)$$

Suppose that $\lim_{Z \rightarrow \infty} g(Z) - f(Z) < 0$. Using (47), this implies that

$$\lim_{Z \rightarrow \infty} g'(Z) - f'(Z) = \infty \quad (48)$$

which is inconsistent with $\lim_{Z \rightarrow \infty} g(Z) - f(Z) < 0$.

In the remainder of the proof, we take as given $\lim_{Z \rightarrow \infty} g(Z) - f(Z) \geq 0$. For $Z > 0$,

$$g(Z) - f(Z) = g(0) - f(0) + \int_0^Z (g'(s) - f'(s))ds \quad (49)$$

Using (46),

$$g(Z) - f(Z) = g(0) - f(0) + \int_0^Z \left((-g(s) + f(s))\frac{s}{\sigma} - \frac{\bar{e}}{\sigma}\varphi\left(\frac{s-\bar{e}}{\sigma}\right)\left(1 - \Phi\left(\frac{s}{\sigma}\right)\right) \right) ds \quad (50)$$

In addition, with $\lim_{Z \rightarrow \infty} g(Z) - f(Z) \geq 0$ in the case under consideration, the fact that $\lim_{Z \rightarrow -\infty} g(Z) - f(Z) > 0$ (since f is positive if and only if $Z > \hat{Z}$, while g is positive for any Z), and the fact that the functions f and g intersect at most once (this is proven in Proof 1 below) imply that $g(Z) - f(Z) \geq 0$ for any Z . This result and (46) imply that $g'(Z) - f'(Z)$

is negative for any $Z > 0$, so that, for any given $Z > 0$,

$$\operatorname{argmax}_{s \in [0, Z]} \{-g(Z) + f(Z)\} = Z \quad (51)$$

Using this result and (50), we get

$$g(Z) - f(Z) < g(0) - f(0) + (-g(Z) + f(Z)) \frac{Z}{\sigma} - \int_0^Z \frac{\bar{e}}{\sigma} \varphi\left(\frac{s - \bar{e}}{\sigma}\right) \left(1 - \Phi\left(\frac{s}{\sigma}\right)\right) ds \quad (52)$$

Since the last term on the right-hand-side of (52) is negative, we get

$$g(Z) - f(Z) < \frac{g(0) - f(0)}{1 + \frac{Z}{\sigma}} \quad (53)$$

Then, for finite $g(0) - f(0)$ and finite σ ,

$$\lim_{Z \rightarrow \infty} \frac{g(0) - f(0)}{1 + \frac{Z}{\sigma}} = 0 \quad (54)$$

It follows from (53) that

$$\lim_{Z \rightarrow \infty} g(Z) - f(Z) = 0 \quad (55)$$

and the proof is complete.

Proof of Proposition 2:

First, we must show that the expression in (39) is smaller than the expression in (35) if and only if $Z > \hat{Z}$. The first term on the right-hand-side of (39) is equal to the only term on the right-hand-side of (35). In addition, because of the definition of \hat{Z} , the second term on the right-hand-side of (39) is negative if and only if $Z > \hat{Z}$. The first part of the proof is complete.

Second, Lemmas 3 and 4 imply that

$$\lim_{Z \rightarrow \infty} \frac{dAR(\sigma)}{d\sigma} = 0 \quad (56)$$

The second part of the proof is complete.

Third, (35) implies that

$$\lim_{Z \rightarrow \infty} \frac{\partial AR(\sigma)}{\partial \sigma} = 0 \quad (57)$$

Using (35) and (39) gives

$$\lim_{Z \rightarrow \infty} \frac{\frac{dAR(\sigma)}{d\sigma}}{\frac{\partial AR(\sigma)}{\partial \sigma}} = \lim_{Z \rightarrow \infty} b \left(1 - \frac{1 - \Phi\left(\frac{Z(\sigma)}{\sigma}\right)}{\Phi\left(\frac{Z(\sigma)}{\sigma}\right) - \Phi\left(\frac{Z(\sigma) - \bar{e}}{\sigma}\right)} \frac{\varphi\left(\frac{Z(\sigma) - \bar{e}}{\sigma}\right) - \varphi\left(\frac{Z(\sigma)}{\sigma}\right)}{\varphi\left(\frac{Z(\sigma)}{\sigma}\right)} \right) = 0 \quad (58)$$

where the last equality is because of (55). The third part of the proof is complete.

Proof of Corollary 2:

As in the Proof of Lemma 1, the incentive constraint with a given Z is

$$\int_Z^\infty b(\pi - Z) \psi(\pi | e = \bar{e}) d\pi - \int_Z^\infty b(\pi - Z) \psi(\pi | e = 0) d\pi = C \quad (59)$$

As $Z \rightarrow \infty$, this equation is satisfied if and only if $C \rightarrow 0$.

Proof 1:

If two continuous functions intersect at least twice, then the difference between their first derivatives at the intersection points alternate signs. Therefore, if the difference between the first derivatives of two continuous functions at their intersection point(s) always has the same sign, then these two functions only intersect once, i.e., the intersection point is unique. We

denote by \bar{Z} any Z such that $f(Z) = g(Z)$. That is,

$$\frac{\varphi\left(\frac{\bar{Z}-\bar{e}}{\sigma}\right) - \varphi\left(\frac{\bar{Z}}{\sigma}\right)}{\varphi\left(\frac{\bar{Z}}{\sigma}\right)} = \frac{\Phi\left(\frac{\bar{Z}}{\sigma}\right) - \Phi\left(\frac{\bar{Z}-\bar{e}}{\sigma}\right)}{1 - \Phi\left(\frac{\bar{Z}}{\sigma}\right)} \quad (60)$$

Equivalently,

$$\left(\varphi\left(\frac{\bar{Z}-\bar{e}}{\sigma}\right) - \varphi\left(\frac{\bar{Z}}{\sigma}\right)\right)\left(1 - \Phi\left(\frac{\bar{Z}}{\sigma}\right)\right) = \varphi\left(\frac{\bar{Z}}{\sigma}\right)\left(\Phi\left(\frac{\bar{Z}}{\sigma}\right) - \Phi\left(\frac{\bar{Z}-\bar{e}}{\sigma}\right)\right) \quad (61)$$

We now compare the first derivatives of the left-hand-side and right-hand-side of (16), as derived in (43) and (45), respectively. At any given Z , (43) is larger than (45) if and only if

$$\left(-\frac{Z-\bar{e}}{\sigma}\varphi\left(\frac{Z-\bar{e}}{\sigma}\right) + \frac{Z}{\sigma}\varphi\left(\frac{Z}{\sigma}\right)\right)\left(1 - \Phi\left(\frac{Z}{\sigma}\right)\right) > -\frac{Z}{\sigma}\varphi\left(\frac{Z}{\sigma}\right)\left(\Phi\left(\frac{Z}{\sigma}\right) - \Phi\left(\frac{Z-\bar{e}}{\sigma}\right)\right) \quad (62)$$

At $Z = \bar{Z}$, (43) is larger than (45) if and only if

$$\frac{\varphi\left(\frac{\bar{Z}-\bar{e}}{\sigma}\right) - \varphi\left(\frac{\bar{Z}}{\sigma}\right)}{\varphi\left(\frac{\bar{Z}}{\sigma}\right)} - \frac{\bar{e}}{\bar{Z}} \frac{\varphi\left(\frac{\bar{Z}-\bar{e}}{\sigma}\right)}{\varphi\left(\frac{\bar{Z}}{\sigma}\right)} < \frac{\Phi\left(\frac{\bar{Z}}{\sigma}\right) - \Phi\left(\frac{\bar{Z}-\bar{e}}{\sigma}\right)}{1 - \Phi\left(\frac{\bar{Z}}{\sigma}\right)} \quad (63)$$

Since $\frac{\bar{e}}{\bar{Z}} \frac{\varphi\left(\frac{\bar{Z}-\bar{e}}{\sigma}\right)}{\varphi\left(\frac{\bar{Z}}{\sigma}\right)} > 0$ for any $\bar{Z} > 0$, and because of (60), we know that (43) is always larger than (45) at any \bar{Z} . We have shown that the first derivative of the left-hand-side of (61) is always larger than the first derivative of the right-hand-side of (61) at a given \bar{Z} which satisfies (60), which implies that the function on the left-hand-side of (61) and the one on the right-hand-side intersect only once, so that \bar{Z} is unique.